

## POWER AND SAMPLE SIZE CALCULATIONS FOR MULTIVARIATE LINEAR MODELS WITH RANDOM EXPLANATORY VARIABLES

GWOWEN SHIEH

NATIONAL CHIAO TUNG UNIVERSITY

This article considers the problem of power and sample size calculations for normal outcomes within the framework of multivariate linear models. The emphasis is placed on the practical situation that not only the values of response variables for each subject are just available after the observations are made, but also the levels of explanatory variables cannot be predetermined before data collection. Using analytic justification, it is shown that the proposed methods extend the existing approaches to accommodate the extra variability and arbitrary configurations of the explanatory variables. The major modification involves the noncentrality parameters associated with the  $F$  approximations to the transformations of Wilks likelihood ratio, Pillai trace and Hotelling-Lawley trace statistics. A treatment of multivariate analysis of covariance models is employed to demonstrate the distinct features of the proposed extension. Monte Carlo simulation studies are conducted to assess the accuracy using a child's intellectual development model. The results update and expand upon current work in the literature.

Key words: MANCOVA, MANOVA, multivariate regression, noncentral  $F$  distribution.

### 1. Introduction

This paper studies the multivariate linear model that provides a basic and convenient framework for repeated measures and longitudinal researches. The general setup of the multivariate linear model encompasses many common statistical models as special cases. When all explanatory variables are quantitative or continuous covariates, the models are called multivariate regression models. When the design matrix contains only indicator variables taking values of zero or one, the models are called multivariate analysis of variance models; while the design matrix contains both continuous covariates and indicator variables, the class of models is called multivariate analysis of covariance (MANCOVA) models. Methods for analyzing data from repeated measures designs have recently received considerable attention in the literature; see Keselman, Algina, and Kowalchuk (2001) for a comprehensive review and their references for related discussions. Traditionally, the values of the explanatory variables in these models are treated as fixed and known and the only variability in the model pertains exclusively to the response variables. These multivariate studies are referred to as fixed (conditional) models. The results would be specific to the particular values of the explanatory variables that are observed or preset by the researcher. However, it is quite common in many multivariate linear model applications that the levels of the explanatory variables for each subject cannot be controlled and are available only after the observations are made, such as the quantitative or continuous covariates in the multivariate regression and multivariate analysis of covariance models. These models are usually referred to as random (unconditional) models. Sampson (1974) studied the problem with normal explanatory variables for both the univariate multiple regression model and the multivariate regression model. Also, Mendoza and Stafford (2001) emphasized the important differences between the random and fixed models when performing confidence

The author wishes to thank the associate editor and the referees for comments which improve the paper considerably. This research was partially supported by a grant from the Natural Science Council of Taiwan.

Requests for reprints should be sent to: Gwown Shieh, Department of Management Science, National Chiao Tung University, Hsinchu, Taiwan 30050, R.O.C., E-mail: gwshieh@mail.nctu.edu.tw

interval computation, power and sample size calculations for the squared multiple correlation coefficient.

To calculate power and sample size in the context of multivariate linear models, Muller and Peterson (1984) and O'Brien and Shieh (1992) suggested several promising approaches using non-central  $F$  approximations to the functions of multivariate test statistics. The methods of Muller and Peterson (1984) seem to have been widely accepted, see Keselman (1998), Muller, LaVange, Ramey, and Ramey (1992), O'Brien and Muller (1993), Rencher (1998, Section 4.4) and Timm (2002, Section 4.16). In contrast, the work of O'Brien and Shieh (1992) is less well known. Nevertheless, it was pointed out in O'Brien and Muller (1993) that the methods of Muller and Peterson (1984) are not invariant to sample size and thus the noncentrality can not be defined. In general, O'Brien and Shieh's (1992) approaches appear to be more accurate than the competing methods of Muller and Peterson (1984) according to their numerical examination. Thus, the formulas of O'Brien and Shieh (1992) are of great potential use and should be properly recognized. However, the existing approaches of Muller and Peterson (1984) and O'Brien and Shieh (1992) do not take into account the variability of explanatory variables. Therefore, these methods are only applicable to fixed models. As pointed out in Sampson (1974), the random explanatory variables do not affect the analysis if the analysis is performed conditionally. However, the resulting power functions for random models are fundamentally different and usually more complicated than that of fixed models. Furthermore, Timm (2002, Section 4.17) cautioned the unknown effect of treating random explanatory variables as fixed in power and sample size calculations. Under the circumstances, it is questionable to apply the existing approaches to multivariate linear models with random explanatory variables without proper recognition and account for the extra variability of explanatory variables.

Glueck and Muller (2003) recently considered the problem of adjusting power for a single baseline covariate with Gaussian distribution in multivariate linear models. They proposed both univariate and multivariate methods based on small sample  $F$  and large sample chi-square approximations of Hotelling-Lawley trace statistic. Although their small sample approximations give accurate results, they are computationally intensive. On the contrary, their multivariate approaches appear to be less accurate. Moreover, a single normal covariate in the model would be inadequate if several key random covariates affect the response variable in important and distinctive ways. A natural generalization to incorporate multiple random covariates and non-normal distributions should be essential to their approach for performing power and sample size calculations in practice. Accordingly, Shieh (2003) suggested several extensions to accommodate arbitrary configurations of the explanatory variables. The noncentral chi-square approximations proposed in Shieh (2003) are based on the Pillai trace and Hotelling-Lawley trace statistics, respectively. In fact, it can be shown that the multivariate method in Glueck and Muller (2003) is a special case of the noncentral chi-square approximation of Hotelling-Lawley trace statistic in Shieh (2003). More importantly, it is found in Shieh (2003) that the noncentral chi-square approximations have a potential problem associated with the control of Type I error rate or the sensitivity to the unbalanced design for the most common research paradigm of multivariate general linear models. Furthermore, they are outperformed by the  $F$  transforms of Hotelling-Lawley trace statistic. Although the focus of Shieh (2003) is on the differences between generalized estimating equations and likelihood-based approaches, the notion of  $F$  transformations of Hotelling-Lawley trace statistic has a clear advantage over their chi-square approximation counterpart. As mentioned above, the available methods are restricted to the  $F$  transformations of Hotelling-Lawley trace statistic. Consequently, the extensions of similar methodology have not been addressed for the other two commonly used test statistics of Wilks likelihood ratio and Pillai trace. This is partially due to the fact that the analytical properties associated with  $F$  approximations of Hotelling-Lawley trace statistic are comparatively easier to justify than those of Wilks likelihood ratio and Pillai trace criteria.

In order to improve the practical usefulness and to overcome the analytical difficulties of power and sample size calculation procedures, this research provides systematic solutions to power analyses and sample size determinations for multivariate linear models with random explanatory variables. The ideas of O'Brien and Shieh (1992) and Shieh (2003) are synthesized to present natural modifications of the existing approaches so that the random setting of explanatory variables is embedded in the procedures. The proposed methodology extends and combines various considerations into one unified framework for the prominent  $F$ -type transformations of Wilks likelihood ratio, Pillai trace and Hotelling-Lawley trace statistics. To demonstrate the versatility of the proposed approaches, detailed discussion is provided for a class of MANCOVA models. Consequently, the aforementioned multivariate linear setting presented in Glueck and Muller (2003) can be viewed as a special case. The distinct features of the proposed approach are the accommodation of an arbitrary number of fixed, random and mixed components of explanatory variables, the flexibility of joint distributions of explanatory variables and the coverage of major multivariate test statistics within the framework of multivariate linear models.

In the next section, the usual multivariate linear model with fixed explanatory variables and the  $F$ -based approaches to power and sample size calculations of O'Brien and Shieh (1992) are described. Section 3 provides the analytical justification and important details of the proposed extension and illustrates the results within the framework of MANCOVA models with both fixed and random explanatory variables. In Section 4, Monte Carlo simulation study is conducted to assess the finite sample adequacy of the proposed methods using a child's intellectual development model. Finally, Section 5 contains some final remarks.

## 2. The $F$ -based Approaches for Fixed Models

Consider the standard multivariate linear model with all the levels of explanatory variables fixed a priori

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \boldsymbol{\varepsilon}, \quad (1)$$

where  $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_N)^T$  is a  $N \times p$  matrix with  $\mathbf{Y}_i$  as the  $p \times 1$  vector of observed sequence of measurements for the  $i$ th subject;  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)^T$  is a  $N \times r$  design matrix with full column rank  $r < N$ , where  $\mathbf{x}_i$  is the  $r \times 1$  vector of explanatory variables associated with the  $i$ th subject;  $\mathbf{B}$  is the  $r \times p$  matrix of unknown regression coefficients; and  $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_N)^T$  is a  $N \times p$  matrix with  $\boldsymbol{\varepsilon}_i$  as the  $p \times 1$  vector of random errors associated with the  $i$ th subject, for  $i = 1, \dots, N$ . The errors  $\boldsymbol{\varepsilon}_i$  are assumed to have independent and identical normal distribution  $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ , where  $\boldsymbol{\Sigma}$  is a  $p \times p$  positive-definite covariance matrix. The general linear hypothesis  $H_0: \mathbf{CBA} = \boldsymbol{\Psi}_0$  is considered here, where  $\mathbf{C}$  is the  $c \times r$  matrix of between-subject contrasts with full row rank  $c \leq r$ , and  $\mathbf{A}$  is the  $p \times a$  matrix of within-subject contrasts with full column rank  $a \leq p$ . The maximum likelihood estimators for  $\mathbf{B}$  and  $\boldsymbol{\Sigma}$  are  $\hat{\mathbf{B}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$  and  $\hat{\boldsymbol{\Sigma}} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})^T(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}/N)$ , respectively. The common statistics for  $H_0: \mathbf{CBA} = \boldsymbol{\Psi}_0$  are obtained from the eigenvalues of  $\mathbf{E}^{-1}\mathbf{H}$ , where

$$\mathbf{E} = (\mathbf{YA} - \mathbf{X}\hat{\mathbf{B}}\mathbf{A})^T(\mathbf{YA} - \mathbf{X}\hat{\mathbf{B}}\mathbf{A}) \text{ and } \mathbf{H} = (\mathbf{C}\hat{\mathbf{B}}\mathbf{A} - \boldsymbol{\Psi}_0)^T[\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T]^{-1}(\mathbf{C}\hat{\mathbf{B}}\mathbf{A} - \boldsymbol{\Psi}_0).$$

With standard results, it can be easily established that  $\mathbf{E}$  and  $\mathbf{H}$  have independent Wishart distributions:

$$\mathbf{E} \sim W_a(N - r, \mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A}) \text{ and } \mathbf{H} \sim W_a(c, \mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A}, \boldsymbol{\Delta}),$$

where  $N - r \geq a$ ,  $c \geq a$  and  $\boldsymbol{\Delta} = (\mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A})^{-1}(\mathbf{CBA} - \boldsymbol{\Psi}_0)^T[\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T]^{-1}(\mathbf{CBA} - \boldsymbol{\Psi}_0)$  is the noncentrality parameter matrix. More specifically, the major test statistics related to the eigenvalues of  $\mathbf{E}^{-1}\mathbf{H}$  for testing the general linear hypothesis are the Wilks likelihood ratio test

statistic  $\Lambda = |\mathbf{E}(\mathbf{E} + \mathbf{H})^{-1}|$ , Pillai trace  $V = \text{tr}[\mathbf{H}(\mathbf{E} + \mathbf{H})^{-1}]$  and Hotelling-Lawley trace  $T = \text{tr}(\mathbf{E}^{-1}\mathbf{H})$ , where  $\text{tr}(\cdot)$  is the trace of a matrix, and Roy's (1953) largest eigenvalue criterion  $\theta = \lambda_1/(1 + \lambda_1)$ , where  $\lambda_1$  is the largest eigenvalue for the nonzero eigenvalues associated with  $\mathbf{E}^{-1}\mathbf{H}$ . It is in general difficult to compute the noncentral distributions of the test statistics since most results involve asymptotic expansion and zonal polynomial function, see Muirhead (1982, Section 10.4). Therefore, more tractable  $F$ -transformations have been proposed for practical use. For Wilks'  $\Lambda = |\mathbf{E}(\mathbf{E} + \mathbf{H})^{-1}|$ , Rao (1951) showed that

$$F_\Lambda = \frac{1 - \Lambda^{1/t}}{\Lambda^{1/t}} \cdot \frac{\text{df}2_\Lambda}{ca} \quad (2)$$

is approximately distributed as an  $F$  random variable with  $ca$  and  $\text{df}2_\Lambda$  degrees of freedom under the null hypothesis, where  $t = 1$  for  $ca \leq 3$  and  $t = \{(c^2a^2 - 4)/(c^2 + a^2 - 5)\}^{1/2}$  for  $ca \geq 4$ , and  $\text{df}2_\Lambda = t\{N - r - (a - c + 1)/2\} - (ca - 2)/2$ . Regarding Pillai trace  $V = \text{tr}[\mathbf{H}(\mathbf{E} + \mathbf{H})^{-1}]$ , an approximate  $F$ -statistic of Pillai (1956) is given by

$$F_V = \frac{V}{s - V} \cdot \frac{\text{df}2_V}{ca} \quad (3)$$

Under the null hypothesis,  $F_V$  has an approximate  $F$  distribution with  $ca$  and  $\text{df}2_V$  degrees of freedom, where  $s = \text{minimum}(c, a)$  and  $\text{df}2_V = s(N - r + s - a)$ . Two common transformations of Hotelling-Lawley trace  $T = \text{tr}(\mathbf{E}^{-1}\mathbf{H})$  are considered here. Pillai and Samson (1959) proposed

$$F_{T1} = T \cdot \frac{\text{df}2_{T1}}{sca}, \quad (4)$$

where  $\text{df}2_{T1} = s(N - r - a - 1) + 2$ . Moreover, McKeon (1974) presented

$$F_{T2} = T \cdot \frac{\text{df}2_{T2}}{hca}, \quad (5)$$

where  $\text{df}2_{T2} = (ca + 2)g + 4$ ,  $g = \{(N - r)^2 - (N - r)(2a + 3) + a(a + 3)\}/\{(N - r)(c + a + 1) - (c + 2a + a^2 - 1)\}$ , and  $h = (\text{df}2_{T2} - 2)/(N - r - a - 1)$ . Under the null hypothesis, both  $F_{T1}$  and  $F_{T2}$  are compared to an  $F$  distribution with numerator degrees of freedom  $ca$ , and denominator degrees of freedom  $\text{df}2_{T1}$  and  $\text{df}2_{T2}$ , respectively. Since no accurate approximation using a single  $F$  transformation is available for Roy's largest eigenvalue criterion, it will not be considered here. In summary, these  $F$ -type tests defined in (2)–(5) are carried out by rejecting  $H_0$  if the  $F$  value is greater than  $F_{ca, \text{df}2, \alpha}$ , where  $F_{ca, \text{df}2, \alpha}$  is the upper  $\alpha$  percentage point of the central  $F$ -distribution  $F(ca, \text{df}2)$  and  $\text{df}2$  represents the corresponding denominator degrees of freedom ( $\text{df}2_\Lambda, \text{df}2_V, \text{df}2_{T1}, \text{df}2_{T2}$ ).

For the case under the alternative hypothesis, it is more difficult to determine the distributions of the test statistics  $\Lambda$ ,  $V$  and  $T$  due to their complexity. Among various approximations to the noncentral distribution of the test statistics, Muller and Peterson (1984) proposed non-central  $F$  approximations to the  $F$ -type statistics (2)–(4). Along the same line of power and sample size calculations within the framework of fixed multivariate linear models, O'Brien and Shieh (1992) considered slightly simplified approximations to the four  $F$ -type statistics (2)–(5). Suppose that  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)^T$  has  $m$  ( $< N$ ) distinct components  $\mathbf{x}_{uj}$  with corresponding proportions  $\pi_j$ ,  $j = 1, \dots, m$ , where  $\sum_{j=1}^m \pi_j = 1$ . Then  $\mathbf{X}^T\mathbf{X}$  can be expressed as  $\mathbf{X}^T\mathbf{X} = \mathbf{N}\mathbf{K}$ , where  $\mathbf{K} = \sum_{j=1}^m \pi_j \mathbf{X}_{uj} \mathbf{X}_{uj}^T$ . Hence,

$$(\mathbf{CBA} - \Psi_0)^T [\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1} \mathbf{C}^T]^{-1} (\mathbf{CBA} - \Psi_0) = \mathbf{N}\mathbf{D},$$

where  $\mathbf{D} = (\mathbf{CBA} - \boldsymbol{\Psi}_0)^T(\mathbf{CK}^{-1}\mathbf{C}^T)^{-1}(\mathbf{CBA} - \boldsymbol{\Psi}_0)$ . It follows that the distribution of  $\mathbf{H}$  can be expressed as  $\mathbf{H} \sim W_a(c, \mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A}, N\bar{\boldsymbol{\Lambda}})$ , where  $\bar{\boldsymbol{\Lambda}} = (\mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A})^{-1}\mathbf{D}$ . Essentially, O'Brien and Shieh (1992) considered the following approximations

$$F_\Lambda \sim F(ca, df2_\Lambda, \delta_\Lambda), \quad F_V \sim F(ca, df2_V, \delta_V),$$

$$F_{T1} \sim F(ca, df2_{T1}, \delta_{T1}) \quad \text{and} \quad F_{T2} \sim F(ca, df2_{T2}, \delta_{T2}), \tag{6}$$

where  $\delta_\Lambda = N \cdot t(\Lambda_{OS}^{-1/t} - 1)$ ,  $\delta_V = N \cdot sV_{OS}/(s - V_{OS})$ ,  $\delta_{T1} = \delta_{T2} = N \cdot T_{OS}$ , with  $\Lambda_{OS} = |(\mathbf{I}_a + \bar{\boldsymbol{\Lambda}})^{-1}|$ ,  $V_{OS} = \text{tr}[\bar{\boldsymbol{\Lambda}}(\mathbf{I}_a + \bar{\boldsymbol{\Lambda}})^{-1}]$ ,  $T_{OS} = \text{tr}(\bar{\boldsymbol{\Lambda}})$  and  $\mathbf{I}_a$  is the  $a \times a$  identity matrix.

### 3. The Proposed Methods

Suppose that the explanatory variables  $\{\mathbf{X}_i^* = \mathbf{x}_i^*, i = 1, \dots, N\}$  follow a distribution  $f(\mathbf{X}_i^*)$  with finite moments, which can be discrete, continuous or mixed such as the interaction between a random continuous covariate and a fixed discrete predictor. The form of  $f(\mathbf{X}_i^*)$  is assumed to be dependent on none of the unknown parameters  $\mathbf{B}$  and  $\boldsymbol{\Sigma}$ . Thus, the notation of  $\mathbf{x}_i$  and  $\mathbf{X}$  as observed values for the multivariate linear model (1) in the previous section are replaced with corresponding  $\mathbf{X}_i^*$  and  $\mathbf{X}^*$  as random variables here, where  $\mathbf{X}^* = (\mathbf{X}_1^*, \dots, \mathbf{X}_N^*)^T$ . Accordingly,  $\hat{\mathbf{B}}, \mathbf{H}$  and  $\mathbf{E}$  are expressed as  $\hat{\mathbf{B}}^* = (\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{X}^{*T}\mathbf{Y}$ ,  $\mathbf{H}^* = (\mathbf{C}\hat{\mathbf{B}}^*\mathbf{A} - \boldsymbol{\Psi}_0)^T[\mathbf{C}(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{C}^T]^{-1}(\mathbf{C}\hat{\mathbf{B}}^*\mathbf{A} - \boldsymbol{\Psi}_0)$  and  $\mathbf{E}^* = (\mathbf{YA} - \mathbf{X}^*\hat{\mathbf{B}}^*\mathbf{A})^T(\mathbf{YA} - \mathbf{X}^*\hat{\mathbf{B}}^*\mathbf{A})$ , respectively. It follows from the standard asymptotic result that  $\mathbf{X}^{*T}\mathbf{X}^*/N = \sum_{i=1}^N \mathbf{X}_i^*\mathbf{X}_i^{*T}/N$  converges in probability to  $\mathbf{K}^*$ , where  $\mathbf{K}^* = E_{\mathbf{X}^*}[\mathbf{X}_i^*\mathbf{X}_i^{*T}]$  and  $E_{\mathbf{X}^*}[\cdot]$  denotes the expectation taken with respect to the distribution of  $\mathbf{X}_i^*$ . Under local alternatives to  $H_0$  of the form  $H_1: \mathbf{CBA} = \boldsymbol{\Psi}_0 + \boldsymbol{\Xi}/N^{1/2}$  for the constant matrix  $\boldsymbol{\Xi}$ , it follows from the application of Slutsky's Theorem that  $\mathbf{H}^*$  converges in distribution to the Wishart distribution  $W_a(c, \mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A}, \boldsymbol{\Lambda}^*)$ , where  $\mathbf{H}^* = (\mathbf{C}\hat{\mathbf{B}}^*\mathbf{A} - \boldsymbol{\Psi}_0)^T[\mathbf{C}(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{C}^T]^{-1}(\mathbf{C}\hat{\mathbf{B}}^*\mathbf{A} - \boldsymbol{\Psi}_0)$  and  $\boldsymbol{\Lambda}^* = (\mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A})^{-1}\boldsymbol{\Xi}^T(\mathbf{CK}^{*-1}\mathbf{C}^T)^{-1}\boldsymbol{\Xi}$ . For the purpose of relating asymptotic power function calculations to the fixed values of  $\mathbf{B}$  and  $\boldsymbol{\Sigma}$  in terms of the local alternatives to  $H_0$ , the following operational and asymptotically equivalent Wishart distribution for the distribution of  $\mathbf{H}^*$  is considered  $\mathbf{H}^* \sim W_a(c, \mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A}, N\bar{\boldsymbol{\Lambda}}^*)$ , where  $\bar{\boldsymbol{\Lambda}}^* = (\mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A})^{-1}\mathbf{D}^*$  and  $\mathbf{D}^* = (\mathbf{CBA} - \boldsymbol{\Psi}_0)^T(\mathbf{CK}^{*-1}\mathbf{C}^T)^{-1}(\mathbf{CBA} - \boldsymbol{\Psi}_0)$ . It can be shown that the distribution of  $\mathbf{E}^*$  is the Wishart distribution  $W_a(N - r, \mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A})$  under both hypotheses for either cases of fixed and random models. Therefore,  $\mathbf{E}^{*-1}\mathbf{H}^* = (\mathbf{E}^*/N)^{-1}(\mathbf{H}^*/N)$  converges in probability to  $\bar{\boldsymbol{\Lambda}}^*$ .

Let  $(F_\Lambda^*, F_V^*, F_{T1}^*, F_{T2}^*)$  represent  $(F_\Lambda, F_V, F_{T1}, F_{T2})$  in which the matrices  $\mathbf{E}$  and  $\mathbf{H}$  are replaced by their counterparts  $\mathbf{E}^*$  and  $\mathbf{H}^*$ , respectively. Likewise, the statistics  $\Lambda, V$  and  $T$  are written as  $\Lambda^*, V^*$  and  $T^*$ , respectively. Then these test statistics depend on the eigenvalues of  $\mathbf{E}^{*-1}\mathbf{H}^*$ , or equivalently the roots  $\phi^*$  in  $|\mathbf{H}^* - \phi^* \mathbf{E}^*| = 0$ . Again, under the notion of large sample approximation, the matrices  $\mathbf{E}^*$  and  $\mathbf{H}^*$  are replaced by their (approximate) expected values  $E[\mathbf{E}^*] = (N - r)\mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A}$  and  $E[\mathbf{H}^*] = c\mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A} + N\mathbf{D}^*$ . This leads to a population version of  $|\mathbf{H}^* - \phi^* \mathbf{E}^*| = 0$ , namely,

$$0 = |c\mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A} + N\mathbf{D}^* - \gamma^*(N - r)\mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A}| = |\mathbf{D}^* - \gamma^*\mathbf{A}^T\boldsymbol{\Sigma}\mathbf{A}| \quad \text{for large } N,$$

where  $\gamma^*$  denote the eigenvalues of  $\bar{\boldsymbol{\Lambda}}^*$ . Consequently, the noncentrality parameters of the statistics  $(F_\Lambda^*, F_V^*, F_{T1}^*, F_{T2}^*)$  are functions of the eigenvalues  $\gamma^*$  of  $\bar{\boldsymbol{\Lambda}}^*$ .

In order to provide a unified formulation of the approximate noncentrality parameter for the  $F$ -type statistic as in O'Brien and Shieh (1992), we extract the factor  $N$  from the  $F$ -type statistic by evaluating the limiting value of  $(ca \cdot F^*/N)$ , namely,

$$f^* = \lim_{N \rightarrow \infty} \left( \frac{ca \cdot F^*}{N} \right).$$

Then, the result is equated to the asymptotic form of the expected value of  $(ca \cdot F^*/N)$ . It can be shown that

$$E \left[ \frac{ca \cdot F^*}{N} \right] = \frac{df2 - 2}{df2} \cdot \frac{ca + \delta^*}{N} \doteq \frac{\delta^*}{N} \quad \text{for large } N,$$

where  $\delta^*$  is the noncentrality parameter of the  $F^*$  statistic. It follows that the noncentrality parameter  $\delta^*$  can be approximated by  $\delta^* = N \cdot f^*$ , where  $f^*$  is the effect size. Hence, the distributions of the  $F$ -type statistics  $(F_\Lambda^*, F_V^*, F_{T1}^*, F_{T2}^*)$  can be approximated as follows:

$$\begin{aligned} F_\Lambda^* &\sim F(ca, df2_\Lambda, \delta_\Lambda^*), F_V^* \sim F(ca, df2_V, \delta_V^*), \\ F_{T1}^* &\sim F(ca, df2_{T1}, \delta_{T1}^*) \text{ and } F_{T2}^* \sim F(ca, df2_{T2}, \delta_{T2}^*), \end{aligned} \tag{7}$$

where the corresponding noncentrality parameters are of the form  $\delta_\Lambda^* = N \cdot t(\Lambda^{*-1/t} - 1)$ ,  $\delta_\Lambda^* = N \cdot sV^*/(s - V^*)$ ,  $\delta_{T1}^* = \delta_{T2}^* = N \cdot T^*$ , with  $\Lambda^* = |(\mathbf{I}_a + \bar{\Lambda}^*)^{-1}|$ ,  $V^* = \text{tr}[\bar{\Lambda}^*(\mathbf{I}_a + \bar{\Lambda}^*)^{-1}]$ , and  $T^* = \text{tr}(\bar{\Lambda}^*)$ . Accordingly, the proposed approximate statistical power achieved for testing hypothesis  $H_0: \mathbf{CBA} = \Psi_0$  with specified significance level  $\alpha$  against the alternative  $H_1: \mathbf{CBA} \neq \Psi_0$  is the probability

$$P\{F(ca, df2, \delta^*) > F_{ca,df2,\alpha}\}, \tag{8}$$

where  $df2 = (df2_\Lambda, df2_V, df2_{T1}, df2_{T2})$  and  $\delta^* = (\delta_\Lambda^*, \delta_V^*, \delta_{T1}^*, \delta_{T2}^*)$  for the approximate  $F$  statistics  $(F_\Lambda^*, F_V^*, F_{T1}^*, F_{T2}^*)$ .

It is important to note that the proposed approximations given in (7) resemble the noncentral  $F$  approximate counterparts of O'Brien and Shieh (1992) given in (6) for fixed models. Conceptually, the resulting formulas and computations for both situations are identical if the probabilities for discrete distribution of  $\mathbf{X}_i^*$  are viewed as the weights for distinct configuration of  $\mathbf{x}_i^*$ . With this conceptual and straightforward modification of notation, the proposed methods apply to the usual multivariate linear models with fixed levels of explanatory variables as well. In general, numerical integration may be needed to carry out the expectation of  $\mathbf{K}^*$  in  $\bar{\Lambda}^*$  for continuous or complex explanatory variables. Note that the proposed approximations for  $F_{T1}^*$  and  $F_{T2}^*$  here are identical to those described in Shieh (2003) based on heuristic extensions. However, not only is the analytical justification presented here more in-depth, the systematic generalization also covers more  $F$  transformations ( $F_\Lambda^*$  and  $F_V^*$ ) than Shieh (2003). The approximate power functions defined in (8) can be inverted to calculate the sample size needed to test hypothesis  $H_0: \mathbf{CBA} = \Psi_0$  against the alternative  $H_1: \mathbf{CBA} \neq \Psi_0$  in order to attain the specified power  $1 - \beta$  for the chosen significance level  $\alpha$ , parameter values  $\mathbf{B}$  and  $\Sigma$ , and probability distribution  $f(\mathbf{X}_i^*)$ . However, it usually involves an iterative process to find the solution because both  $F(ca, df2, \delta^*)$  and  $F_{ca,df2,\alpha}$  depend on the sample size  $N$ .

To demonstrate the general setup of the proposed approaches, the class of MANCOVA models is considered. Let the design matrix  $\mathbf{X}^*$  in the multivariate linear model be of the form  $\mathbf{X}_{FG} = [\mathbf{F} \ \mathbf{G}]$  where  $\mathbf{F}$  represents the  $N \times r_F$  matrix of fixed explanatory variables and  $\mathbf{G}$  denotes the  $N \times r_G$  matrix of random explanatory variables with  $r = r_F + r_G$ . As illustrated in Section 2 for the fixed setting, suppose that  $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_N)^T$  has  $m$  finite number of distinct values  $\mathbf{f}_{uj}$  with corresponding proportions  $\pi_j$ ,  $j = 1, \dots, m$ , where  $\sum_{j=1}^m \pi_j = 1$ . Moreover, the random component  $\mathbf{G} = (\mathbf{G}_1, \dots, \mathbf{G}_N)^T$  is composed of random vectors  $\mathbf{G}_i$  where  $\mathbf{G}_i$  follows a distribution  $f(\mathbf{G}_i)$  with finite moments,  $i = 1, \dots, N$ . Specifically, assume that the mean vector  $E_{G_i}[\mathbf{G}_i] = \mu_G$  and covariance matrix  $\text{Var}[\mathbf{G}_i] = \Sigma_G$ . It follows that  $\mathbf{X}_{FG}^T \mathbf{X}_{FG} / N$  is adequately approximated by  $\mathbf{K}_{FG}^*$  under large sample consideration, where

$$\mathbf{K}_{FG}^* = \begin{bmatrix} \mathbf{K}_{11}^* & \mathbf{K}_{12}^* \\ \mathbf{K}_{12}^{*T} & \mathbf{K}_{22}^* \end{bmatrix},$$

$\mathbf{K}_{11}^* = \sum_{j=1}^m \pi_j \mathbf{f}_{uj} \mathbf{f}_{uj}^T$ ,  $\mathbf{K}_{12}^* = \sum_{j=1}^m \pi_j \mathbf{f}_{uj} \boldsymbol{\mu}_G^T$  and  $\mathbf{K}_{22}^* = \boldsymbol{\mu}_G \boldsymbol{\mu}_G^T + \boldsymbol{\Sigma}_G$ . As a result, with  $\mathbf{X}^* = \mathbf{X}_{FG}$  and  $\mathbf{K}^* = \mathbf{K}_{FG}^*$ , the proposed general formulas for the four statistics given in (7) can immediately be applied to performing power and sample size calculations in the MANCOVA models. It is noteworthy that the proposed approaches do not require a full specification of the distribution form for the explanatory variables; only the mean vector and covariance matrix of the explanatory variables are needed. If the investigator is unable or unwilling to specify a distribution for  $\mathbf{G}$ , the observed  $n \times r_G$  matrix  $\mathbf{G}_{obs}$  from the pilot study can be employed as an empirical approximation to the underlying distribution  $f(\mathbf{G}_i)$ . Specifically, one can proceed to approximate the distribution of  $\mathbf{G}_i$  with  $p(\mathbf{G}_i = \mathbf{G}_{uj}) = 1/n$ , where  $\mathbf{G}_{uj}$ ,  $j = 1, \dots, n$ , are the  $n$  distinct columns of  $\mathbf{G}_{obs}^T$ . Hence, the proposed approach is still applicable in such situations.

For the special case of a single random covariate ( $r_G = 1$ ) with mean 0 and variance  $\sigma_G^2$ , it is obvious that  $\mathbf{K}_{12}^*$  and  $\mathbf{K}_{22}^*$  reduce to the  $r_F \times 1$  null vector  $\mathbf{0}$  and  $\sigma_G^2$ , respectively. Hence,  $\mathbf{K}_{FG}^*$  is simplified to  $\mathbf{K}_{FG1}^*$ , where

$$\mathbf{K}_{FG1}^* = \begin{bmatrix} \mathbf{K}_{11}^* & \mathbf{0} \\ \mathbf{0}^T & \sigma_G^2 \end{bmatrix}.$$

Essentially, the proposed formulas can be readily established under this particular specification. For this particular model, Glueck and Muller (2003) suggested both small sample  $F$  and large sample  $\chi^2$  approximations to estimate the power of Hotelling-Lawley trace statistic. Although their small sample approximations are very accurate, they require the integration of conditional power with the density function of noncentrality. The computation appears to be considerably more complicated than the proposed  $F$  approximations. In addition, their large sample noncentral  $\chi^2$  approximation has the identical noncentrality as the  $F_{T2}$  approximation given in (7) with  $\delta_{T2}^* = N \cdot \text{tr}(\bar{\Delta}_{FG1}^*)$  and  $\bar{\Delta}_{FG1}^* = (\mathbf{A}^T \boldsymbol{\Sigma} \mathbf{A})^{-1} (\mathbf{CBA} - \boldsymbol{\Psi}_0)^T (\mathbf{CK}_{FG1}^{*-1} \mathbf{C}^T)^{-1} (\mathbf{CBA} - \boldsymbol{\Psi}_0)$ . It is interesting to note that the illustration and argument for the derivation of noncentrality parameter  $\bar{\Delta}_{FG1}^*$  presented here are much simpler than those in Lemma 3 and Theorem 3 of Glueck and Muller (2003). More importantly, it is well known that the chi-square approximation is too liberal in controlling the type I error under finite-sample assessment. According to the findings in Shieh (2003), the phenomenon continues to exist in the case of random explanatory variables. Hence, its practical use is problematic.

#### 4. Simulation Study

In this section, Monte Carlo simulation study is conducted to evaluate the finite sample property of the proposed approaches. For illustration, the model formulation of the child’s intellectual development example considered in Muller et al. (1992, Section 3) is exploited as the base for the numerical examinations. The example involves a longitudinal study of a child’s intellectual performance as a function of the mother’s estimated verbal intelligence. With child IQ measurements at 12, 24, and 36 months ( $p = 3$ ), and with intercept, linear, quadratic, and cubic trends in the mother’s standardized IQ (MSIQ) as explanatory variables ( $r = 4$ ), this yields

$$\mathbf{Y}_i = \begin{bmatrix} \text{IQ}_{12} \\ \text{IQ}_{24} \\ \text{IQ}_{36} \end{bmatrix}, \mathbf{X}_i^* = \begin{bmatrix} 1 \\ \text{MSIQ} \\ \text{MSIQ}^2 \\ \text{MSIQ}^3 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} \beta_{1.12} & \beta_{1.24} & \beta_{1.36} \\ \beta_{L.12} & \beta_{L.24} & \beta_{L.36} \\ \beta_{Q.12} & \beta_{Q.24} & \beta_{Q.36} \\ \beta_{C.12} & \beta_{C.24} & \beta_{C.36} \end{bmatrix},$$

where  $\text{IQ}_t$  is the child’s IQ measurement at time  $t$ ,  $\beta_{1,t}$  is the intercept, while  $\beta_{L,t}$ ,  $\beta_{Q,t}$ , and  $\beta_{C,t}$  are the corresponding coefficients of linear, quadratic, and cubic values of MSIQ for time  $t = 12$ ,

24, and 36, respectively. In an attempt to demonstrate the usefulness of the proposed methods for studies with random explanatory variables, it is assumed that the mother’s standardized IQ has a standard normal distribution,  $MSIQ \sim N(0, 1)$ . It follows that

$$\mathbf{K}^* = E_{\mathbf{X}^*}[\mathbf{X}_i^* \mathbf{X}_i^{*T}] = \begin{bmatrix} 1 & m_1 & m_2 & m_3 \\ m_1 & m_2 & m_3 & m_4 \\ m_2 & m_3 & m_4 & m_5 \\ m_3 & m_4 & m_5 & m_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 3 \\ 1 & 0 & 3 & 0 \\ 0 & 3 & 0 & 15 \end{bmatrix},$$

where  $m_i$  is the  $i$ th moment of a standard normal distribution. According to the description in Muller et al. (1992), the model parameter matrices are set as

$$\mathbf{B} = \begin{bmatrix} 114.46 & 104.66 & 98.83 \\ 2.88 & 8.77 & 10.67 \\ -0.71 & -0.90 & -1.30 \\ -0.21 & -0.54 & -0.72 \end{bmatrix} \text{ and } \mathbf{\Sigma} = \begin{bmatrix} 218.48 & 83.66 & 72.19 \\ 83.66 & 251.92 & 158.60 \\ 72.19 & 158.60 & 244.58 \end{bmatrix}.$$

The hypothesized relationship between mother and child competence of interest corresponds to a test of the time  $\times$  mother’s IQ interaction  $H_0: \mathbf{CBA} = 0$  with the between-subject and within-subject contrast matrices ( $c = 3$  and  $a = 2$ )

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{A} = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}$$

respectively.

With the specifications described above, the simulation study is conducted in three steps. First, the effect sizes and estimates of sample sizes required for testing the specified hypothesis with significance level  $\alpha = 0.05$  and power levels (0.80, 0.90) are calculated. The resulting effect sizes and sample sizes correspond to the proposed statistics and are presented in Table 1.

TABLE 1.  
Calculated sample sizes and estimates of actual power at specified sample size for the child development model

	$F_{\Lambda}^*$		$F_{\nabla}^*$		$F_{T1}^*$		$F_{T2}^*$	
MSIQ $N(0, 1)$								
Effect size	0.1288		0.1248		0.1328		0.1328	
N for power 0.80 and 0.90	110	139	113	143	106	135	108	137
Estimated $\alpha$	0.0523	0.0488	0.0508	0.0485	0.0530	0.0499	0.0523	0.0492
Nominal power at $N_{\Lambda}$	0.8042	0.9013	0.7896	0.8905	0.8181	0.9111	0.8112	0.9074
Estimated power	0.8024	0.8996	0.7961	0.8980	0.8070	0.9017	0.8051	0.9001
MSIQ standardized Gamma(5, 2)								
Effect size	0.1216		0.1184		0.1248		0.1248	
N for power 0.80 and 0.90	116	147	119	151	113	143	115	145
Estimated $\alpha$	0.0506	0.0468	0.0491	0.0455	0.0514	0.0481	0.0506	0.0470
Nominal power at $N_{\Lambda}$	0.8030	0.9012	0.7907	0.8922	0.8148	0.9096	0.8082	0.9060
Estimated power	0.7872	0.8819	0.7815	0.8790	0.7917	0.8844	0.7889	0.8833
MSIQ standardized Gamma(10, 2)								
Effect size	0.1220		0.1186		0.1254		0.1254	
N for power 0.80 and 0.90	115	146	119	151	112	143	114	144
Estimated $\alpha$	0.0502	0.0467	0.0496	0.0452	0.0510	0.0478	0.0497	0.0464
Nominal power at $N_{\Lambda}$	0.8004	0.9000	0.7873	0.8904	0.8128	0.9089	0.8062	0.9052
Estimated power	0.7885	0.8920	0.7833	0.8889	0.7937	0.8953	0.7902	0.8931



In the second step, the sample sizes in the following simulations are unified by choosing the sample size estimate associated with  $F_{\Lambda}^*$  in step 1, say  $N_{\Lambda}$ , as the benchmark. Under the null hypothesis  $H_0: \mathbf{CBA} = \mathbf{0}$  with the given sample sizes and model configurations, the estimates of actual Type I error rate for the nominal significance level  $\alpha = 0.05$  are computed through Monte Carlo simulation using 10,000 replicate data sets. For each replicate,  $N_{\Lambda}$  values of MSIQ are generated from standard normal distribution. In turn, these values determine the mean functions for generating  $N_{\Lambda}$  children's intellectual performance outcomes. The computations are conducted by comparing the simulated values of test statistics with their corresponding critical values  $F_{6, df2, 0.05}$ . The estimated  $\alpha$  is the proportion of the 10,000 replicates whose test statistic values exceed the critical value.

The third and last step studies the power approximations under the alternative hypothesis  $H_1: \mathbf{CBA} \neq \mathbf{0}$ . For a fair comparison among these approaches, the nominal powers at  $N_{\Lambda}$  are recalculated for all four competing methods. The SAS/IML (SAS Institute, 2003) program used to perform power calculations of the proposed methods is provided in the Appendix. As expected, the nominal powers of  $F_{\Lambda}^*$  are almost identical to 0.80 or 0.90, while those of the other three tests ( $F_V^*$ ,  $F_{T1}^*$ ,  $F_{T2}^*$ ) slightly deviate from the values of 0.80 or 0.90. As in the previous step, the estimates of true power are computed using 10,000 replicate data sets. The adequacy of the sample size formula is determined by the difference between the estimated and nominal values of power. All calculations are performed using programs written with SAS/IML (2003).

As suggested by a referee, the vector of explanatory variables composed of powers of a standardized Gamma variable for MSIQ is also considered. Specifically,  $\mathbf{X}_i^* = [1 \ Z \ Z^2 \ Z^3]^T$ , where  $Z = (X - E[X]) / [\text{Var}(X)]^{1/2}$  and  $X$  has a Gamma ( $g_1, g_2$ ) distribution with the shape and scale parameters  $g_1$  and  $g_2$ , respectively. For illustrative purpose, the parameters ( $g_1, g_2$ ) are set as (5, 2) and (10, 2). It can be shown that the matrix  $\mathbf{K}^*$  for the standardized Gamma(5, 2) and standardized Gamma(10, 2) distributions of MSIQ are

$$\mathbf{K}^* = \begin{bmatrix} 1 & 0 & 1 & 0.8944 \\ 0 & 1 & 0.8944 & 4.2 \\ 1 & 0.8944 & 4.2 & 11.0909 \\ 0.8944 & 4.2 & 11.0909 & 45.8 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 1 & 0.6325 \\ 0 & 1 & 0.6325 & 3.6 \\ 1 & 0.6325 & 3.6 & 7.0835 \\ 0.6325 & 3.6 & 7.0835 & 29.2 \end{bmatrix},$$

respectively. Ultimately, the simulation process was repeated with these two standardized Gamma settings and the corresponding empirical results are also presented in Table 1.

It can be seen from the results summarized in Table 1 that first, the computed sample sizes allow comparison of relative efficiencies of the four approaches. It is interesting to note that the ordering of sample size is consistently  $F_V^* > F_{\Lambda}^* > F_{T2}^* > F_{T1}^*$  in all cases considered here. However, it appears that the differences in magnitude are small. Among the three MSIQ distributions, the two standardized Gamma distributions incur slightly larger sample size estimates than the standard normal situation for all four  $F$  transformations. Next, the resulting values of estimated  $\alpha$  are surprisingly accurate for all four statistics throughout Table 1. However, it should be noted that the errors between nominal powers and estimated powers associated with the skewed Gamma distributions are relatively larger than those of the symmetric normal distribution. This situation is more prominent for the standardized Gamma(5, 2) distribution due to the outsized moments. In general, the computed power approximations for the competing methods maintain a close agreement between the estimated power and nominal power. Specifically, the method  $F_{T1}^*$  gives the largest errors among the four competing formulas. Thus, for the two methods based on  $F$  transforms of Hotelling-Lawley trace statistic, the  $F_{T2}^*$  approach constantly provides more accurate results. This finding is in accordance with those of Shieh (2003). More importantly, both the proposed extensions  $F_{\Lambda}^*$  and  $F_V^*$  have excellent performance of achieving the nominal levels in all standard normal and standardized Gamma settings for MSIQ. Overall, the accuracy

of the proposed approaches increases slightly with the sample size, and varies with the structure of the model parameters. Nevertheless, the results are sufficiently accurate for most purposes. In practical situations, it is a difficult task to assess the robustness of the proposed approaches for the underlying distribution of explanatory variables. For the cases of complicated and unbalanced explanatory variables distributions, it is advisable to consider a range of design variations to provide guidance about the sample sizes required for the study.

## 5. Conclusions

For the power and sample size calculations within the framework of multivariate linear models, this article provides updated description of existing methods of O'Brien and Shieh (1992). However, their approaches only apply to the fixed models whose results would be specific to the particular realization of the explanatory variables. As in the child's development example, the explanatory variables typically can not be fixed in advance and induce additional variability to the paradigm. Similar applications are frequently encountered in behavioral and social studies. For practical purpose, the work of O'Brien and Shieh (1992) and Shieh (2003) are exploited thoroughly to obtain useful methods and promising results. The simple structure of  $F$  transformations permits computational simplifications that are explicitly recognized in the statistical procedures for analysis of variance. The essence of the proposed approaches is the modification of the noncentrality for the  $F$ -type statistics to accommodate the characteristics of random explanatory variables. For implementation, the only difference is that the designation of particular values for design matrix amounts to the specification of the joint distribution of explanatory variables. In fact, the general formulation allows fixed, random and mixed components of explanatory variables. The numerical assessment suggests that the discrepancies between the estimated and nominal levels of type I error rate and power seemed completely acceptable, given the many unknowns in study planning. The proposed methods are efficient, accurate and more simplified than the other approximations in similar studies. According to these findings, it is concluded that the proposed methods expand the current literature and should be valuable in many applications.

## Appendix

### *SAS/IML program for performing power calculations of the proposed methods*

```

PROC IML;
ALPHA=0.05;N=110;
B={114.46 104.66 98.83, 2.88 8.77 10.67, -0.71 -0.90 -1.30, -0.21 -0.54 -0.72};
SIGMA={218.48 83.66 72.19, 83.66 251.92 158.60, 72.19 158.60 244.58};
KSTAR={1 0 1 0,0 1 0 3,1 0 3 0,0 3 0 15};
CMAT={0 1 0 0, 0 0 1 0, 0 0 0 1};
AMAT=((-1,0,1)/SQRT(2))||({1,-2,1}/SQRT(6));
R=NROW(B);P=NCOL(B);C=NROW(CMAT);A=NCOL(AMAT);S=MIN(A,C);CA=C#A;
SMALLN=N-R;
EPM=(AMAT`*SIGMA*AMAT);
HPM=(CMAT*B*AMAT)`*INV(CMAT*INV(KSTAR)*CMAT)`*(CMAT*B*AMAT);
EIHPM=INV(EPM)*HPM;
*L;
IF CA < 4 THEN T=1;ELSE T=SQRT((CA##2-4)/(C##2+A##2-5));
DF2L=T#(SMALLN-(A-C+1)/2)-(CA-2)/2;
FCRITL=FINV(1-ALPHA,CA,DF2L,0);

```

```

LPM=DET(INV(I(A)+EIHPM));EFSL=T#(LPM##(-1/T)-1);
POWERL=1-PROBF(FCRITL,CA,DF2L,N#EFSL);
*V;
DF2V=S#(SMALLN+S-A);
FCRITV=FINV(1-ALPHA,CA,DF2V,0);
VPM=TRACE(EIHPM*INV(I(A)+EIHPM));EFSV=S#VPM/(S-VPM);
POWERV=1-PROBF(FCRITV,CA,DF2V,N#EFSV);
*T1;
DF2T1=S#(SMALLN-A-1)+2;
FCRITT1=FINV(1-ALPHA,CA,DF2T1,0);
T1PM=TRACE(EIHPM);EFST1=T1PM;
POWER1=1-PROBF(FCRITT1,CA,DF2T1,N#EFST1);
*T2;
G=(SMALLN##2-SMALLN#(2#A+3)+A#(A+3))/
  (SMALLN#(C+A+1)-(C+2#A+A##2-1));
DF2T2=4+(CA+2)#G;
FCRITT2=FINV(1-ALPHA,CA,DF2T2,0);
T2PM=TRACE(EIHPM);EFST2=T2PM;
POWER2=1-PROBF(FCRITT2,CA,DF2T2,N#EFST2);
PRINT N[FORMAT=6.0],
  EFSL[FORMAT=7.4] EFSV[FORMAT=7.4] EFST1[FORMAT=7.4]
  EFST2[FORMAT=7.4],
  POWERL[FORMAT=7.4] POWERV[FORMAT=7.4] POWER1[FORMAT=7.4]
  POWER2[FORMAT=7.4];
QUIT;

```

## References

- Glueck, D.H., & Muller, K.E. (2003). Adjusting power for a baseline covariate in linear models. *Statistics in Medicine*, 22, 2535–2551.
- Keselman, H.J. (1998). Testing treatment effects in repeated measures designs: An update for psychophysiological researchers. *Psychophysiology*, 35, 470–478.
- Keselman, H.J., Algina, J., & Kowalchuk, R.K. (2001). The analysis of repeated measures designs: a review. *British Journal of Mathematical and Statistical Psychology*, 54, 1–20.
- McKeon, J.J. (1974). F approximations to the distribution of Hotelling's  $T_0^2$ . *Biometrika*, 61, 381–383.
- Mendoza, J.L., & Stafford, K.L. (2001). Confidence interval, power calculation, and sample size estimation for the squared multiple correlation coefficient under the fixed and random regression models: A computer program and useful standard tables. *Educational and Psychological Measurement*, 61, 650–667.
- Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. New York, NY: Wiley.
- Muller, K.E., & Peterson, B.L. (1984). Practical methods for computing power in testing the multivariate general linear hypothesis. *Computational Statistics and Data Analysis*, 2, 143–158.
- Muller, K.E., LaVange, L.M., Ramey, S.L., & Ramey, C.T. (1992). Power calculations for general linear multivariate models including repeated measures applications. *Journal of the American Statistical Association*, 87, 1209–1226.
- O'Brien, R.G., & Muller, K.E. (1993). Unified power analysis for t-tests through multivariate hypotheses. In L.K. Edwards (Ed.), *Applied Analysis of Variance in Behavioral Science* (pp. 297–344). New York, NY: Marcel Dekker.
- O'Brien, R.G., & Shieh, G. (1992). Pragmatic, unifying algorithm gives power probabilities for common F tests of the multivariate general linear hypothesis. In paper presented at the Annual Joint Statistical Meetings of the American Statistical Association, Boston, Massachusetts.
- Pillai, K.C.S. (1956). On the distribution of the largest or the smallest root of a matrix in multivariate analysis. *Biometrika*, 43, 122–127.
- Pillai, K.C.S., & Samson, P. Jr. (1959). On Hotelling's generalization of  $T^2$ . *Biometrika*, 46, 160–168.
- Rao, C.R. (1951). An asymptotic expansion of the distribution of Wilks' criterion. *Bulletin of the International Statistics Institute*, 33, 177–180.
- Rencher, A.C. (1998). *Multivariate Statistical Inference and Applications*. New York, NY: Wiley.
- Roy, S.N. (1953). On a heuristic method of test construction and its use in multivariate analysis. *Annals of Mathematical Statistics*, 24, 220–238.
- Sampson, A.R. (1974). A tale of two regressions. *Journal of the American Statistical Association*, 69, 682–689.
- SAS Institute (2003). *SAS/IML software: usage and reference*. Version 8, Cary, NC.

- Shieh, G. (2003). A comparative study of power and sample size calculations for multivariate general linear models. *Multivariate Behavioral Research*, 38, 285–307.
- Timm, N.H. (2002). *Applied Multivariate Analysis*. New York, NY: Springer.
- Manuscript received 28 MAY 2003*  
*Final version received 23 MAR 2004*